Adaboost

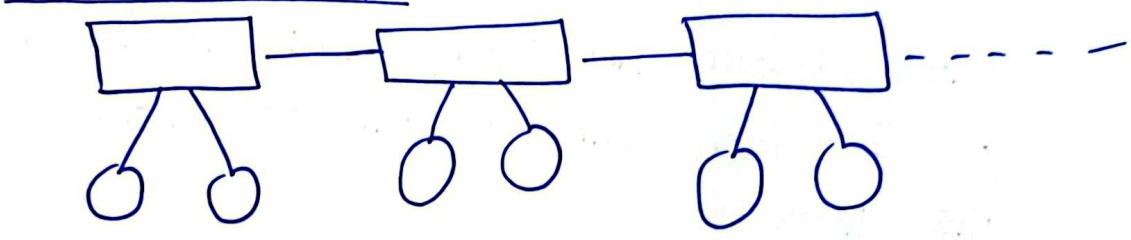
In Adaboost, assign weights to the weak learners

$$f = \alpha_1 M_1 + \alpha_2 M_2 + \alpha_3 M_3 + \ldots\ldots + \alpha_n M_n$$

$\therefore M_1 \ldots M_n$: are decision tree stumps

$\therefore \alpha_1 \ldots \alpha_n$: are weights

Decision Tree Stumps



Decision Tree Stumps mean max depth = 1

↳ weak learner

↳ Underfitting

↳ Train Accuracy ↓ 40% } High Bias

↳ Test Accuracy ↕ 45% } Low Variance

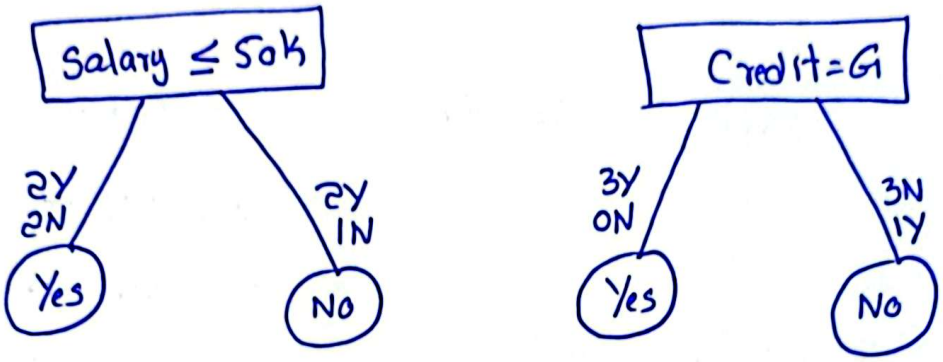
By combining multiple weak learners

↳ Low Bias, High Variance

→ Mathematical Intuition

Salary	Credit	Approval
<=50K	B	No
<=50K	G	Yes
<=50K	G	Yes
>50K	B	No
>50K	G	Yes
>50K	N	Yes
<=50K	N	No

Step:1: Create Decision Tree Stumps



which decision tree stump needs to be selected as the first stump, we will use entropy or gini index

~~Gini(Salary ≤ 50) = 1 -~~

~~Gini(Credit = G) = 1 -~~

$$\text{Gini}(\text{Salary} \leq 50 \text{ and Yes}) = 1 - \left(\frac{2}{4}\right)^2 - \left(\frac{2}{4}\right)^2 = 0.5$$

$$\text{Gini}(\text{Salary} \leq 50 \text{ and No}) = 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 = 0.44$$

$$\text{Gini}(\text{Salary} \leq 50) = \frac{4}{7}(0.5) + \frac{3}{7}(0.44)$$

$$= 0.46$$

$$\text{Gini}(\text{Credit} = G \text{ and Yes}) = 1 - \left(\frac{3}{3}\right)^2 - \left(\frac{0}{3}\right)^2 = 0$$

$$\text{Gini}(\text{Credit} = G \text{ and No}) = 1 - \left(\frac{1}{4}\right)^2 - \left(\frac{3}{4}\right)^2 = 0.375$$

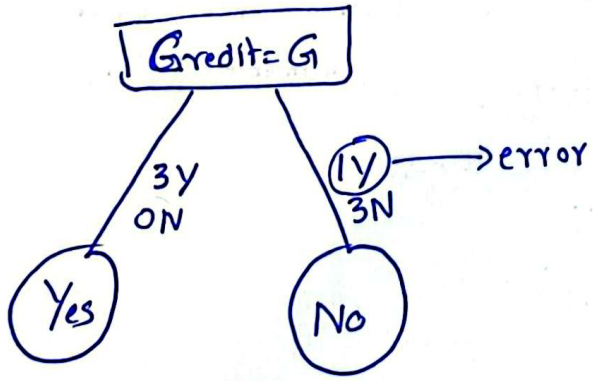
$$\text{Gini}(\text{Credit}) = \frac{3}{7}(0) + \frac{4}{7}(0.375)$$

$$= 0.16$$

Credit = G is selected as first decision stump

Step 2: Sum of the Total Errors and Performance of the stumps

Salary	Credit	Approval	Sample weights	
<=50K	B	No	1/7	
<=50K	G	Yes	1/7	
<=50K	G	Yes	1/7	
>50K	B	No	1/7	
>50K	G	Yes	1/7	
>50K	N	Yes	1/7	← Error
<=50K	N	No	1/7	



① Sum of total errors = $\frac{1}{7}$

② Performance of stump = $\frac{1}{2} \ln\left[\frac{1 - \text{Total Error}}{\text{Total Error}}\right]$

$= \frac{1}{2} \ln\left[\frac{1 - 1/7}{1/7}\right]$

$= \frac{1}{2} \ln[6]$

$= 0.896$

Performance of stump ≈ 0.896

$f = \alpha_1 M_1 + \ldots + \alpha_n M_n$

$\alpha \rightarrow 0.896$

Step 3: Update the weights for correctly and incorrectly classified points

Salary	Credit	Approval	Sample weights	Updated weights
<=50K	B	No	1/7	0.058
<=50K	G	Yes	1/7	0.058
<=50K	G	Yes	1/7	0.058
>50K	B	No	1/7	0.058
>50K	G	Yes	1/7	0.058
>50K	N	Yes	1/7	0.349
<=50K	N	No	1/7	0.058

① decrease the weights for correct classified points

$$= \text{weight} \times e^{-\text{performance of stump}}$$

$$= \frac{1}{7} \times e^{-0.896}$$

$$= 0.058$$

② Update the weight for incorrect classified points

$$= \text{weight} \times e^{\text{performance of stump}}$$

$$= \frac{1}{7} \times e^{0.058}$$

$$= 0.349$$

Step 4: Normalizing weights and Assigning Bins

Salary	Credit	Approval	Updated Weights	Normalized Weights
<=50K	B	No	0.058	0.08
<=50K	G	Yes	0.058	0.08
<=50K	G	Yes	0.058	0.08
>50K	B	No	0.058	0.08
>50K	G	Yes	0.058	0.08
>50K	N	Yes	0.349	0.50
<=50K	N	No	0.058	0.08
			0.697	1

① divide all weights by 0.697

② Model 1 sends only incorrectly classified records to Model 2, will be assigning some bins

Salary	Credit	Approval	Updated Weights	Normalized Weights	Bins
<=50K	B	No	0.058	0.08	0.00 - 0.08
<=50K	G	Yes	0.058	0.08	0.08 - 0.16
<=50K	G	Yes	0.058	0.08	0.16 - 0.24
>50K	B	No	0.058	0.08	0.24 - 0.32
>50K	G	Yes	0.058	0.08	0.32 - 0.40
>50K	N	Yes	0.349	0.50	0.40 - 0.90
<=50K	N	No	0.058	0.08	0.90 - 0.98

Step 5: Select data points to send to next stump

Salary	Credit	Approval	Bins
<=50K	B	No	0 - 0.08
<=50K	G	Yes	0.08 - 0.16
<=50K	G	Yes	0.16 - 0.24
>50K	B	No	0.24 - 0.32
>50K	G	Yes	0.32 - 0.40
>50K	N	Yes	0.40 - 0.90
<=50K	N	No	0.90 - 0.98

① Iterate process by selecting random values between 0 and 1 and check where they fall in the bins

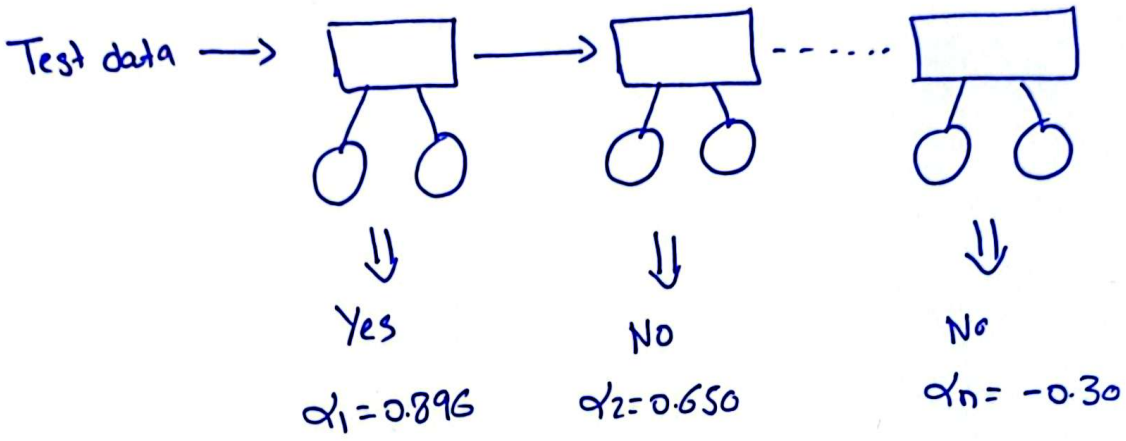
Salary	Credit	Approval	Random Number	
>50K	N	Yes	0.50	←
<=50K	G	Yes	0.10	
>50K	N	Yes	0.60	←
>50K	B	No	0.32	

Step: 6 These records will be sent to next decision tree stump

Step 2 - Step 6 will be repeated

Step 7: Final Prediction

Test data ($\leq 50K, G$)



$$f = \alpha_1 M_1 + \alpha_2 M_2 + \cdots + \alpha_n M_n$$

$$= (0.896)(Yes) + (0.650)(No) - 0.30(No)$$

$$= (0.896)(Yes) + (0.350)(No)$$

→ maximum is Yes

$0.896 > 0.350$

Final output: Yes